

Arquitecturas de Generación de Contenido Colaborativo para sistemas basados en Realidad Aumentada Móvil

Daniel Gallego Vico, Iván Martínez Toro y Joaquín Salvachúa Rodríguez.

Departamento de Ingeniería de Sistemas Telemáticos

Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid

Avenida Complutense nº 30, "Ciudad Universitaria", 28040, Madrid, España

{dgallego, imartinez, jsalvachua}@dit.upm.es

Resumen- La evolución actual de los terminales móviles ha propiciado el surgimiento de un nuevo campo de investigación relacionado con las aplicaciones móviles colaborativas basadas en Realidad Aumentada. Debido a su inmadurez, es necesaria una conceptualización de términos que aclaren un entorno hasta el momento complejo y poco estructurado. Este artículo propone una nueva taxonomía llamada "Pirámide de Generación de Contenido Colaborativo" que clasifica este tipo de aplicaciones en tres niveles: aisladas, sociales y en tiempo real. Dicha clasificación describe las diferentes arquitecturas que se deben tener en cuenta para conseguir sistemas de cada uno de estos niveles, teniendo en cuenta la forma en que el contenido aumentado es generado y cómo se lleva a cabo la colaboración. Por tanto, el principal objetivo es clarificar terminología relativa a este nuevo paradigma, a la vez que se propone un marco para identificar y clasificar futuras investigaciones relativas a este entorno.

Palabras Clave- Realidad Aumentada, Móvil, Colaboración, Arquitecturas, Localización, Generación de contenido, Taxonomía, Experiencia de usuario.

I. INTRODUCCIÓN

En los últimos años, el mundo móvil ha experimentado una evolución increíblemente rápida potenciada aún más gracias al surgimiento de los *smartphones*. La disponibilidad de grandes capacidades técnicas tanto en las infraestructuras de red como en los terminales, unido al esfuerzo puesto en integrar grandes redes sociales como Facebook o Twitter, ha permitido la creación de lo que se ha venido a llamar colaboración móvil [1].

Por otro lado, es importante destacar que recientemente hemos contemplado el renacimiento de la Realidad Aumentada (RA) hasta convertirse en un tema de actualidad [2] gracias en gran medida a la evolución de los terminales móviles. De esta forma, las propiedades que un sistema basado en RA debe cumplir según Azuma et al. [3], están soportadas perfectamente por los actuales dispositivos que presentan potencias de cómputo y capacidades multimedia acordes a los requisitos. En consecuencia, y como se ha visto en los últimos años en el mercado móvil, los desarrolladores han encontrado que la combinación entre las técnicas asociadas al área de la RA y los dispositivos móviles ha propiciado un fértil campo de investigación y desarrollo donde experimentar y crear innovadoras aplicaciones. Además, desde el punto de vista de los usuarios, la inclusión

de la RA en el entorno móvil ha cambiado radicalmente la forma en que interactuamos con el mundo, ya que a través de estos sistemas estamos enriqueciendo la realidad con un contenido aumentado o virtual que nos provee de nuevos tipos de información y modelos de visualización, modificando por tanto la experiencia de usuario final.

Por estas razones, el presente artículo está enfocado en la colaboración móvil basada en RA (una unión consecuencia directa de los avances en ambos campos) que se apoya en teléfonos móviles que no utilizan tecnologías marcadoras, ropa inteligente o cualquier otro tipo de tecnología que usualmente no está presente en un *smartphone*. Se ha elegido este enfoque ya que la aceptación de los teléfonos móviles es un hito ya conseguido hace años, mientras que el de tecnologías más experimentales de *tracking* se encuentra en un estado más exploratorio que real, y por tanto lo expuesto en estas líneas estará dirigido a terminales al alcance de toda la sociedad, que cumplen unos criterios de comodidad, transparencia óptica, precio medio y atractivo acordes al uso de teléfonos móviles tal y como Feiner expone en [4].

Estudiar este tipo de sistemas y las arquitecturas relativas a ellos, así como los diferentes parámetros que se ven involucrados será por tanto el objetivo de este artículo, en el que presentaremos una taxonomía denominada *Pirámide de Generación de Contenido Colaborativo* que mediante tres niveles (Aislado, Social y Real) nos permitirá clasificar las diferentes aplicaciones móviles colaborativas basadas en RA.

Por tanto, gracias a esta taxonomía veremos como analizar y reconocer las ventajas y desventajas de cada una de las diferentes arquitecturas, además de permitirnos clasificar las aplicaciones que actualmente podemos encontrar en el mercado móvil.

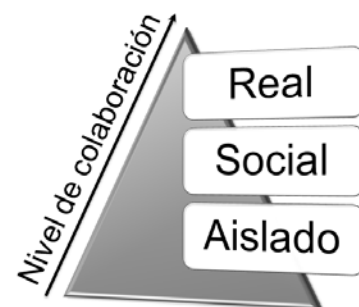


Fig. 1. Pirámide de Generación de Contenido Colaborativo.

II. PIRÁMIDE DE GENERACIÓN DE CONTENIDO COLABORATIVO

La RA es hoy en día un área de investigación y desarrollo muy activa, y debido a su rápido crecimiento y relativa inmadurez sufre la falta de cierta conceptualización. Es por ello que varias de sus interesantes propiedades no son recogidas en ninguna taxonomía existente, como por ejemplo las diferencias entre aplicaciones en cuanto a generación de contenido. Las definiciones, así como las taxonomías, son herramientas esenciales que llevan a la optimización de los procesos de investigación y desarrollo. Una muestra de esta necesidad es el esfuerzo que está llevando a cabo el *World Wide Web Consortium* (W3C) para definir estándares que gestionen los recursos involucrados en la RA, como pueden ser la geolocalización [5] o el acceso a dispositivos como la cámara [6].

Durante el estudio del entorno de la RA y las aplicaciones que coexisten en el mismo, hemos apreciado la existencia de una característica concreta que difiere entre distintas aplicaciones. Esta característica define la manera en que se genera el contenido a ser mostrado en forma de capas de RA. La capacidad de colaboración permitida por cada aplicación está directamente relacionada con esta característica.

Para cubrir la falta de una taxonomía general que represente la característica anteriormente introducida, proponemos una nueva categorización: la Pirámide de Generación de Contenido Colaborativo (ilustrada en la Figura 1). Esta taxonomía clasifica el entorno de las aplicaciones móviles colaborativas basadas en RA tomando como criterio principal la característica explicada anteriormente, pero también tiene en cuenta otros conceptos estrictamente relacionados como la tecnología y la arquitectura necesarias, y el impacto percibido por el usuario final.

Además, la estructura piramidal de la clasificación proviene del hecho de que cada nivel es soportado por la arquitectura y tecnología de los niveles inferiores. Luego, siguiendo estas ideas hemos definido tres niveles diferenciados (Aislado, Social y Real), que serán explicados en detalle en las siguientes secciones aportando ejemplos de aplicaciones reales en Apple Store o en Android Market para validar su uso como herramienta de clasificación.

A. Primer nivel: colaboración aislada

Este nivel incluye cualquier aplicación que utilice una gestión y generación de contenido centralizada. Esto es, cualquier aplicación en la que la información que formará las diferentes capas sobre el mundo real es creada o recopilada por el equipo que soporta la propia aplicación y almacenada y gestionada únicamente en los servidores de la misma. No hay ningún tipo de contribución por parte del usuario final al repositorio de contenidos, siendo el usuario únicamente un consumidor de información, sin participar en la generación ni la mejora de la misma. Como resultado se obtiene que la colaboración en este nivel es mínima.

Sin embargo, este es un nivel que es importante estudiar no solo por la existencia de un gran número de aplicaciones que concuerdan con sus características, sino también por ser la base tecnológica y arquitectural del resto de niveles superiores.



Fig. 2. Arquitectura general del nivel de colaboración aislado.

La arquitectura que soporta este tipo de sistemas es en la mayoría de los casos similar a la mostrada en la Figura 2. La parte principal del esquema es el servidor de la aplicación, que contiene todo el contenido a ser mostrado en forma de capas de información. Denominamos este nivel Aislado debido a que el servidor no se relaciona con otros servicios, siendo por tanto la única fuente de información para la parte cliente. El cliente suele ser un *smartphone* que cuenta con diferentes dispositivos de captura de contexto físico como la cámara, la brújula o el acelerómetro, y que es capaz de acceder a recursos externos como el GPS. Diversas aplicaciones se valen de distintos subconjuntos de estos dispositivos para generar las capas de información de RA.

Las aplicaciones pertenecientes a este nivel están principalmente basadas en información sobre el contexto del usuario final. Utilizan información del entorno del usuario para mejorar los mecanismos de la generación de RA, optimizar su rendimiento [7], y estar alerta de todo lo que ocurre en el entorno para reaccionar de forma acorde al contexto del usuario [8].

Como ejemplo de aplicaciones de este nivel, hemos seleccionado dos: *Nearest Tube* y *Theodolite*, ambas completamente basadas en técnicas de captación de contexto y en las que sólo se utilizan los servidores de la propia aplicación para generar las capas de RA.

B. Segundo nivel: colaboración social

El nivel Social está localizado en la parte media de la pirámide y se refiere a aquellas aplicaciones que presentan capas de información generadas a partir de diferentes fuentes de contenido. No sólo grandes medios como Wikipedia, o la propia aplicación comparten sus contenidos, sino también los propios usuarios finales generan y comparten información a través de redes sociales (Facebook, Twitter, Blogger, etc.) o subiéndola directamente al servidor de la aplicación. Cada usuario es capaz de seleccionar un Punto De Interés (PDI) y adjuntar cierta información al mismo, que será almacenada y mostrada al resto de usuarios. El entorno social de cada usuario colabora de forma dinámica para crear contenido, lo que se traduce en generación de contexto social que acompaña al físico ya obtenido en el nivel anterior.



Fig. 3. Arquitectura general del nivel de colaboración social.

En este caso la arquitectura es más compleja que en el nivel inferior ya que ha de soportar la colaboración entre diferentes participantes, tal y como puede observarse en la Figura 3. El servidor de aplicación tiene que permitir la adición dinámica de contenido, mientras que la parte cliente es similar a la descrita para el nivel Aislado, con la capacidad extra de subir contenido al servidor y de aceptarlo de diversas fuentes.

Se deduce de la descripción anterior que la colaboración móvil en RA comienza en este nivel. El potencial de las redes sociales combinado con el de la RA y las características de los terminales móviles, permiten la creación de aplicaciones verdaderamente colaborativas en las que el usuario final es quien crea la información más interesante. Otros usuarios consumirán dicha información en forma de capas sobre el mundo real. Este nivel también genera información de contexto social que puede ser utilizado para hacer posible un abanico de aplicaciones colaborativas de distintos tipos [9].

La mayoría de las aplicaciones que están siendo desarrolladas y puestas en producción hoy en día se enmarcan dentro de este nivel de la pirámide. Gracias a su creciente importancia y aceptación, nuevas formas de colaboración están siendo habilitadas. De este modo, tenemos ejemplos reales como Layar o Junaio que permiten añadir piezas simples de información (comentarios, valoraciones, etc.) sobre PDIs existentes, o como Wikitude, WhereMark y Sekai Camera que permiten la creación y edición de PDIs.

C. Tercer nivel: colaboración en tiempo real

El nivel superior de la pirámide corresponde a las aplicaciones que comparten capas de información generadas en tiempo real. Específicamente, en este tipo de aplicaciones dos o más usuarios se conectan entre ellos para contribuir en la generación en vivo del contenido aumentado en forma de capas sobre el PDI capturado por uno de los usuarios. El contenido generado puede ser compartido de cara a estar disponible para usos futuros aprovechando la conexión con redes sociales o servidores propios del nivel anterior.



Fig. 4. Arquitectura general del nivel de colaboración en tiempo real.

Para dar soporte a este tipo de aplicaciones es necesario contar con una arquitectura completamente distribuida. En la Figura 4 está representada la abstracción de una posible arquitectura, mostrando la interacción entre usuarios y la generación de contenido sobre el área de análisis de uno de ellos, así como la forma en que se comparte dicho contenido. Las aplicaciones basadas en este nivel requieren toda la tecnología ofrecida por los niveles inferiores y, además, una solución tecnológica para la comunicación en tiempo real.

Este nivel supone un área de investigación con un futuro prometedor. Esto es así debido a que hoy en día no es posible encontrar aplicaciones comerciales capaces de interconectar varios usuarios para generar contenido aumentado en tiempo real de manera colaborativa. Por esta razón, a continuación describimos brevemente dos posibles casos de uso enmarcados en diferentes áreas de conocimiento. El primero, en educación, se basaría en impartir una clase al completo siguiendo un modelo de conexión de 1 a N usuarios, con 1 profesor y N estudiantes recibiendo las enseñanzas a través las pantallas de sus teléfonos móviles y permitiendo además la posibilidad de añadir notas en una capa de RA y compartirlas con el resto de la clase. El segundo caso, esta vez relacionado con la medicina, consiste en conectar a pacientes situados en zonas geográficamente aisladas o de difícil acceso con el médico de familia, o permitir exámenes médicos a distancia realizados por uno o más médicos de manera simultánea y colaborativa.

III. CONCLUSIONES Y TRABAJOS FUTUROS

A lo largo de este artículo hemos propuesto una serie de definiciones con el objetivo de conceptualizar y clasificar el paradigma actual de las aplicaciones móviles basadas en RA, mostrando además las arquitecturas generales que se encuentran detrás de este tipo de sistemas. Por ello, llegados a este punto es necesario analizar diferentes cuestiones importantes relativas a la propuesta de una nueva taxonomía.

En primer lugar, es necesario describir el rol que juega dicha taxonomía frente a otras planteadas anteriormente, que

como se ha comentado ya, eran escasas y poco enfocadas a esta área. Así, debemos comparar la “Pirámide de Generación de Contenido Colaborativo” con la clasificación “*Interaction Techniques and User Interfaces*” descrita en [10]. Específicamente, en dicha aportación se define la colaboración remota basada en RA como la capacidad para integrar múltiples usuarios que poseen diferentes dispositivos y que se encuentran en contextos dispares, incrementando la interacción entre ellos en tiempo real gracias a una mejora de su experiencia de usuario. Si nos detenemos un instante a pensar en esta idea, comprobamos que el tercer nivel de nuestra taxonomía (nivel Real) es la evolución lógica de esta definición propuesta en 2008. Esto es así, ya que nuestra propuesta intenta establecer la creación de un espacio compartido entre usuarios que permita colaborar mediante la generación de contenido virtual sobre un entorno mixto que proporcione un contexto colaborativo avanzado.

Otro punto importante a analizar es cómo ayuda la taxonomía descrita en este artículo a evaluar el entorno de las aplicaciones móviles basadas en RA. Una respuesta rápida es que nos permite clasificarlas de manera adecuada y con unos criterios claros basados en algo fácilmente identificable como es la arquitectura del sistema y los métodos de generación de contenido colaborativo, poniendo orden por tanto en un entorno bastante caótico que ha evolucionado sin una línea clara en los últimos años. Sin embargo, hay una respuesta si cabe más interesante: ofrece a diseñadores y desarrolladores la oportunidad de trabajar en un marco común claramente definido y conocido por todos, que además, ayuda a entender más fácilmente que tipo de aplicación se quiere implementar y qué módulos son necesarios. Así, identificar por ejemplo en el mercado móvil los competidores que se tiene a la hora de desarrollar una nueva aplicación resulta más sencillo.

Por otro lado, y enfocando ahora nuestro análisis en detalles de índole más tecnológica, creemos que existen significativas líneas de estudio que pueden ser explotadas en años venideros, siendo algunas de las más importantes las que detallamos a continuación.

La primera que queremos remarcar está relacionada con uno de los mayores problemas que podemos encontrar cuando se generan capas de información de RA en tiempo real. Concretamente nos referimos al renderizado de gráficos (muchas veces en 3D) en tiempo real, que es necesario para visualizar las capas aumentadas de este tipo de aplicaciones. Estos procesos son pesados en lo que a procesamiento se refiere, y aunque los actuales *smartphones* son dispositivos muy potentes, es evidente que descargar a los clientes móviles de este tipo de operaciones daría mayor libertad para dedicar esos recursos a otros aspectos. Por ello, creemos que iniciativas como OnLive [11] son altamente interesantes, ya que actualmente ofrece un portal de videojuegos que se ejecuta en la nube, siendo necesarios únicamente una conexión de banda ancha y no un ordenador potente para ejecutar dichos juegos, pues el renderizado gráfico se realiza en sus servidores. Luego, aplicar este tipo de técnicas de *cloud computing* [12] al entorno móvil permitiría que no sólo los últimos modelos de *smartphones* pudiesen soportar aplicaciones de los niveles superiores de la pirámide, sino también aquellos menos potentes.

Cambiando de enfoque, y poniendo nuestra atención ahora en la poca estandarización que este entorno posee

actualmente, sería sumamente interesante crear un API o formato común para la definición de PDIs, de manera que una aplicación pudiese integrar PDIs creados en cualquier otra aplicación, evitando así la replicación y redundancia de contenido que actualmente existe entre aplicaciones. Con esto en mente, hemos seguido por un lado lo propuesto en el *W3C Workshop: Augmented Reality on the Web* donde Reynolds et al. [13] propusieron usar Linked Data [14] para definir PDIs en aplicaciones de RA, y por otro lado el proyecto actual de definición de PDIs en el que W3C está trabajando [15].

Por último, y enmarcado en un trabajo futuro relacionado con la propia definición de la taxonomía, sería interesante añadir una nueva dimensión a la misma que tuviese en cuenta el impacto social de usar las aplicaciones de los diferentes niveles, como por ejemplo el número máximo de usuarios que es viable que colaboren a la vez o cómo este tipo de sistemas afectan a la experiencia de usuario final.

REFERENCIAS

- [1] F. Reynolds, “Web 2.0-In Your Hand”. *IEEE Pervasive Computing*, vol. 8, no. 1, pp. 86-88, Jan. 2009.
- [2] S.J. Vaughan-Nichols, “Augmented Reality: No Longer a Novelty?” *Computer*, vol. 42, no. 12, pp. 19-22, Dec. 2009.
- [3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier and B. MacIntyre, “Recent advances in Augmented Reality”. *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34-47, Nov/Dec. 2001.
- [4] S.K. Feiner, “The importance of being mobile: some social consequences of wearable augmented reality systems”. In *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, pp. 145-148, San Francisco, CA, USA, Oct. 1999.
- [5] W3C, “Geolocation API Specification”, 2010. [Online]. Available: <http://www.w3.org/TR/geolocation-API/> [Accessed: March 10, 2011].
- [6] W3C, “HTML Media Capture”, 2010. [Online]. Available: <http://www.w3.org/TR/capture-api/> [Accessed: March 10, 2011].
- [7] W. Lee and W. Woo, “Exploiting Context-Awareness in Augmented Reality Applications”. In *Proceedings of the International Symposium on Ubiquitous Virtual Reality*, pp. 51-54, Gwangju, South Korea, Jul. 2008.
- [8] T. Hofer, W. Schwinger, M. Pichler, G. Leonhartsberger, J. Altmann and W. Retschitzegger, “Context-awareness on mobile devices - the hydrogen approach”. In *Proceedings of the 36th Annual Hawaii International Conference on System Science*, pp. 292-301, Big Island, Hawaii, Jan. 2003.
- [9] E. Prasolova-Forland, M. Divitini and A.E. Lindas, “Supporting Social Awareness with 3D Collaborative Virtual Environments and Mobile Devices: VirasMobile”. In *Proceedings of the Second International Conference on Systems*, pp. 33-38, Sainte-Luce, Martinique, France, April 2007.
- [10] Feng Zhou, H.B.-L. Duh and M. Billinghurst, “Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR”. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 193-202, Cambridge, UK, Sep. 2008.
- [11] OnLive, 2010. [Online]. Available: <http://www.onlive.com/> [Accessed: March 10, 2011].
- [12] X. Luo, “From Augmented Reality to Augmented Computing: A Look at Cloud-Mobile Convergence”. In *Proceedings of the International Symposium on Ubiquitous Virtual Reality*, 29-32, Guangju, Korea, Jul. 2009.
- [13] V. Reynolds, M. Hausenblas, A. Polleres, M. Hauswirth and V. Hegde, “Exploiting Linked Open Data for Mobile Augmented Reality”. In *Proceedings of the W3C Workshop: Augmented Reality on the Web*, Barcelona, Spain, Jun. 2010.
- [14] Linked Data, Connect Distributed Data Across the Web, 2010. [Online]. Available: <http://linkeddata.org/> [Accessed: March 10, 2011].
- [15] W3C, Points of Interest (POI) Working Group, 2010. [Online]. Available: <http://www.w3.org/2010/POI/> [Accessed: March 10, 2011].